```
                    GenCore version 4.5

             Copyright (c) 1993 - 2000 Compugen Ltd.


OM protein - protein search, using sw model

Run on:          March  1, 2001, 15:43:16 ; Search time 299.73 Seconds
                                            (without alignments)
                                            260.436 Million cell updates/sec

Title:           US-09-331-631A-1
Perfect score:   3542
Sequence:        1 MAINTSNLCSLLFLLSLFLL.........SPRSTKQQQPLVSILDFVGF 666

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters:          374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    SPTREMBL_15:*
              1: sp_archea   :*
              2: sp_bacteria:*
              3: sp_fungi:*
              4: sp_human:*
              5: sp_invertebrate:*
              6: sp_mammal:*
              7: sp_mhc:*
              8: sp_organelle:*
              9: sp_phage:*
             10: sp_plant:*
             11: sp_rodent:*
             12: sp_virus:*
             13: sp_vertebrate:*
             14: sp_unclassified:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	3542	100.0	666	10	Q9SPL5	macadamia i
2	3412	96.3	666	10	Q9SPL4	macadamia i
3	3227	91.1	625	10	Q9SPL3	macadamia i
4	1314.5	37.1	593	10	Q9SEW4	juglans reg
5	1198.5	33.8	810	10	Q9ZWI3	cucurbita m
6	1025	28.9	525	10	Q43358	theobroma c
7	898	25.4	582	10	Q43358	theobroma c
8	864.5	24.4	637	10	O03865	zea mays (m
9	842	23.8	613	10	O03678	hordeum vul
10	812	22.9	544	10	Q9M3X6	pisum sativ
11	802.5	22.7	489	10	O22120	glycine max
12	787	22.2	545	10	Q9SPI1	glycine max
13	769.5	21.7	483	10	Q41674	vicia narbo
14	752	21.2	448	10	Q49927	pisum sativ
15	749	21.1	486	10	Q40873	picea glauc
16	747	21.1	417	10	O9LUJ7	arabidopsis
17	741.5	20.9	438	10	O22221	glycine max
18	741	20.9	450	10	Q40844	picea glauc
19	731.5	20.7	463	10	Q41677	vicia narbo

20	705.5	19.9	518	10	Q9M3X8	lens culina
21	639.5	18.1	461	10	Q9ZRG9	oryza sativ
22	551.5	15.6	232	10	Q9SQ49	rulingia ma
23	549	15.5	430	10	Q41115	phaseolus v
24	547.5	15.5	414	10	Q41727	zamia furfu
25	545	15.4	430	10	Q43633	phaseolus v
26	539	15.2	421	10	Q43632	phaseolus v
27	537	15.2	239	10	Q9SQ50	abroma augu
28	534.5	15.1	236	10	Q9SQ41	theobroma c
29	533.5	15.1	236	10	Q9SQ47	theobroma s
30	532.5	15.0	236	10	Q9SQ36	theobroma g
31	532.5	15.0	236	10	Q9SQ34	theobroma m
32	531.5	15.0	236	10	Q9SQ40	theobroma c
33	531.5	15.0	236	10	Q9SQ37	theobroma g
34	528.5	14.9	236	10	Q9SQ48	theobroma m
35	527.5	14.9	540	10	O03866	zea mays (m
36	527	14.9	239	10	Q9SQ35	theobroma c
37	527	14.9	239	10	O5s7V9	theobroma c
38	525	14.8	239	10	Q9SQ42	theobroma b
39	524	14.8	238	10	Q9SQ32	theobroma s
40	523.5	14.8	236	10	Q9SQ33	theobroma m
41	523	14.8	239	10	Q9SQ44	theobroma c
42	521	14.7	235	10	Q9SQ38	theobroma g
43	519.5	14.7	236	10	Q9SQ46	theobroma a
44	518	14.6	239	10	Q9SQ43	theobroma b
45	516	14.6	423	10	Q43617	phaseolus l

ALIGNMENTS

```
RESULT    1
Q9SPL5

ID   Q9SPL5            PRELIMINARY;      PRT;     666 AA.
AC   Q9SPL5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN PRECURSOR.
GN   AMP2.
OS   Macadamia integrifolia (Macadamia nut).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX   NCBI_TaxID=60698;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=NUT KERNEL;
RA   Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT   "A family of antimicrobial peptides is produced by processing of a 7S
RT   globulin protein in Macadamia integrifolia kernels.";
RL   Plant J. 0:0-0(1999).
RL   EMBL; AF161883; AAD54244.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
SQ   SEQUENCE  666 AA;  78217 MW;  C752B884B2DF0224 CRC64;
```

```
Query Match            100.0%;   Score 3542;  DB 10;  Length 666;
Best Local Similarity  100.0%;   Pred. No. 2.3e-257;
Matches  666;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

Qy       1 MAINTSNLCSLLFLLSLFLLSTTVSLAESEFDRQEYEECKRQCMQLETSGQMRRCVSQCD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MAINTSNLCSLLFLLSLFLLSTTVSLAESEFDRQEYEECKRQCMQLETSGQMRRCVSQCD 60

Qy      61 KRFEEDIDWSKYDNQEDPQTECQCQCRRCRQQESGPRQQQYCQRRCKEICEEEEEYNRQR 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KRFEEDIDWSKYDNQEDPQTECQCQCRRCRQQESGPRQQQYCQRRCKEICEEEEEYNRQR 120

Qy     121 DPQQQYEQCQKHCQRRETEPRHMQTCQQRCERRYEKEKRKQQKRYEEQOREDEEKYEERM 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 DPQQQYEQCQKHCQRRETEPRHMQTCQQRCERRYEKEKRKQQKRYEEQOREDEEKYEERM 180
```

```
Qy  181 KEEDNKRDPQQQREYEDCRRRCEQQEQPRQQHQCQLRCREQQEQQHGIRGRGGDMNPQRGGSGRY 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 KEEDNKRDPQQQREYEDCRRRCEQQEQPRQQHQCQLRCREQQEQQHGIRGRGGDMNPQRGGSGRY 240

Qy  241 EEGEEEQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EEGEEEQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNA 300

Qy  301 FVLPTHLDADAILLVIGGRGALKMIHIHDNRESYNLECGDVIRIPAGTTFYLINRDNNERL 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 FVLPTHLDADAILLVIGGRGALKMIHIHDNRESYNLECGDVIRIPAGTTFYLINRDNNERL 360

Qy  361 HIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLRGVFGQQREGV 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 HIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLRGVFGQQREGV 420

Qy  421 IIRASQEQIRELTRDDSESRHHHIRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYR 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 IIRASQEQIRELTRDDSESRHHHIRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYR 480

Qy  481 QLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKR 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 QLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKR 540

Qy  541 HEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVVSSGNENLLLFAFGINAQNNHENFLAGR 600
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 HEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVVSSGNENLLLFAFGINAQNNHENFLAGR 600

Qy  601 ERNVLQOIEPQAMELAFAAPRKEVEESFNSQDQSIFFPGPRQHQOQQSPRSTKQQQPLVSI 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ERNVLQOIEPQAMELAFAAPRKEVEESFNSQDQSIFFPGPRQHQOQQSPRSTKQQQPLVSI 660

Qy  661 LDFVGF 666
        ||||||
Db  661 LDFVGF 666
```

RESULT 2

Q9SPL4

ID Q9SPL4   PRELIMINARY;   PRT;   666 AA.
AC Q9SPL4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR.
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL;
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7s
RT globulin protein in Macadamia integrifolia.";
RL Plant J. 0:0-0(1999).
DR EMBL; AF161884; AAD54245.1; -.
DR HSSP; P02853; 2PHL.
DR INTERPRO; IPR001113; -.
DR PFAM; PF00546; Seedstore_7s; 1.
SQ SEQUENCE 666 AA; 78243 MW; 0ECA22F8710F8A7B CRC64;

Query Match 96.38;
Best Local Similarity 96.18;
Matches 640; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

Score 3412; DB 10; Length 666;
Pred. No. 1.3e-247;

```
Qy  1 MAINTSNLCSLLFLLSLFLLSTTVSLAESEFDRQEYEBCKRQCMQLETSGQMRRCVSQCD 60
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1 MAINTSNLCSLLFLLSLFLLSTTVSLAESEFDRQEYEBCKRQCMQLETSGQMRRCVSQCD 60
```

RESULT 3

Q9SPL3

```
Qy   61 KRFEEDIDWSKYDNQEDPQTECQQCQQCRRCRQQESGPRQQYCQRRCKEICEEEEEYNRQR 120
        ||||||||||||||||: |||| |::|||||||||||||||||||||||||||||||||||
Db   61 KRFEEDIDWSKYDNQDDPQTDCQQCQQCRRCRQQESGPRQQYCQRRCKEICEEEEEYNRQR 120

Qy  121 DPQQQYEQCQCQKHCQRRETEPRHMQTCCQQCQERRYEKEKRRQQKRYEEQQQREDEEKYEERM 180
        ||||||||||||| :|: |:||||||||||||||||||||||||||||||||||||||||||
Db  121 DPQQQYEQCQCQRHETEPRHMQTCCQQCQERRYEKEKRRQQKRYEEQQQREDEEKYEERM 180

Qy  181 KEEDNKRDPQQQREYEDCRRRCEQQEQPRQQHQCQLRCREQQEQQHGIRGRGGDMNPQRGGSGRY 240
        ||||||||||||||||||||||||||||||||||||||||||||::|||||||||||||||||
Db  181 KEEDNKRDPQQQREYEDCRRRCEQQEQPRQQHQCQLRCREQQEQQHGIDLINPQRGGSGRY 240

Qy  241 EEGEEEQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNA 300
        |||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 EEGEEKQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPNA 300

Qy  301 FVLPTHLDADAILLVIGGRGALKMIHIHDNRESYNLECGDVIRIPAGTTFYLINRDNNERL 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 FVLPTHLDADAILLVTGGRGALKMIHRDNRESYNLECGDVIRIPAGTTFYLINRDNNERL 360

Qy  361 HIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLRGVFGQQREGV 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 HIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQAERLRGVLGQQREGV 420

Qy  421 IIRASQEQIRELTRDDSESRHHHIRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYR 480
        |||:||||||||||||||:||||||||||||||||||||||||||||||||||||||||
Db  421 IISASQEQIRELTRDDSESRRWHIRRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYR 480

Qy  481 QLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKR 540
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  481 QLQDMDVSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKR 540

Qy  541 HEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVVSSGNENLLLFAFGINAQNNHENFLAGR 600
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  541 HEEEEDVHYEQVKARLSKREAIVVPGHPVVFVVSSGNENLLLFAFGINAQNNHENFLAGR 600

Qy  601 ERNVLQOIEPQAMELAFAAPRKEVEESFNSQDQSIFFPGPRQHQOQQSPRSTKQQQPLVSI 660
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  601 ERNVLQOIEPQAMELAFAAPRKEVEELFNSQDESIFFPGPRQHQOQQSSRSTKQQQPLVSI 660

Qy  661 LDFVGF 666
        ||||||
Db  661 LDFVGF 666
```

ID Q9SPL3   PRELIMINARY;   PRT;   625 AA.
AC Q9SPL3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR (FRAGMENT).
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL;
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7s
RT globulin protein in Macadamia integrifolia.";
RL Plant J. 0:0-0(1999).
DR EMBL; AF161885; AAD54246.1; -.
DR HSSP; P02853; 2PHL.
DR INTERPRO; IPR001113; -.
DR PFAM; PF00546; Seedstore_7s; 1.
FT NON_TER   1
SQ SEQUENCE 625 AA; 73586 MW; 415808A89D370296 CRC64;

RESULT 4

```
ID   Q9SEW4
AC   Q9SEW4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).
OS   Juglans regia (English walnut).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fagales; Juglandaceae; Juglans.
OX   NCBI_TaxID=51240;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. SUNLAND; TISSUE=SOMATIC EMBRYO LINE;
RA   Teuber S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.;
RT   "Identification and cloning of a cDNA encoding a vicilin-like protein,
RT   Jug r 2, from English walnut kernel (Juglans regia): a major food
RT   allergen.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF066055; AAF18269.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR001113; -.
```

```
Qy    42 QCMQLETSGQMRRCVSQCDKRFEEDIDWSKYDNQEDPQTECQQCQRRCRCQQESGPRQQQY 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 QCMQLETSGQMRRCVSQCDKRFEEDIDWSKYDNQEDPQTECQQCQRRCRCQQESGPRQQQY 60

Qy   102 CQRRCKEICEEEEYNRQRDPQQQYEQCQKHCQRRETEPRHMQTCQQRCERRYEKEKRKQ 161
          ||||||||||||  ||||||||||||||||||||||||||||||||||||  |  ||
Db    61 CQRRCKEICEEEEYNRQRDPQQQYEQCQKCQRRETEPRHMQICQQRCERRYEKEKRKQ 120

Qy   162 QKRYEEQQQREDEEKYEERMKEEDNKRDPQQOREYEDCRRRCEQQEPRQQHQCQLRCRCEQQQR 221
          ||||||||  |||||||||||||||||||||||||||||||||||::|||:||:|||
Db   121 QKRYEEQQQREDEEKYEERMKEGDNKRDPQQOREYEDCRRHCEQQEPRLQYQCQRRCQEQQR 180

Qy   222 QHGRGGDMMNPQRGGSGRYEEGGEEQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSK 281
          ||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 QHGRGGDLMNPQRGGSGRYEEGEGEEKQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSK 240

Qy   282 LLRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHHDNRESYNLECGDVI 341
          |||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||
Db   241 LLRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHRDNRESYNLECGDVI 300

Qy   342 RIPAGTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAA 401
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 RIPAGTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAA 360

Qy   402 LNTQTEKLRGVFGGQQOREGVIIRASQEQIRELTRDDSESRHWHIRRGGESSRGPYNLFNKR 461
          ||||||||||||||||||||||||||||||||||||:||||||||||||||||||||||
Db   361 LNTQTERLRGVLGQQOREGVIIRASQEQIRELTRDDSESRRWHIRRGGESSRGPYNLFNKR 420

Qy   462 PLYSNKYCGQAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADV 521
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 PLYSNKYCGQAYEVKPEDYRQLQDMDLSVFIANITQGSMMGPFFNTRSTKVVVVASGEADV 480

Qy   522 EMACPHLSGRHGGRGGGKRHEEEEDVHYEQVRARLSKREAIVLAGHPVVFVSSGNENLL 581
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 EMACPHLSGRHGGRGGGKRHEEEEVHYEQVRARLSKREAIVVLAGHPVVFVSSGNENLL 540

Qy   582 LFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAAAPRKEYEESFNSQDQSIFFPGR 641
          ||||||||||||||||||||||||||||||||||||||||||:|||::||||||||||
Db   541 LFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAASRKEVEELFNSQDESIFFPGPR 600

Qy   642 QHQQQSPRSTKQQQPLVSILDFVGF 666
          |||||||||||||||||||||||||
Db   601 QHQQQSPRSTKQQQPLVSILDFVGF 625
```

RESULT 5

```
SQ   SEQUENCE   593 AA;  69990 MW;  9BA127E19B18C0D8 CRC64;

DR   PFAM; PF00546; Seedstore_7s; 1.
NON_TER   1
FT   1

Query Match        37.1%;  Score 1314.5;  DB 10;  Length 593;
Best Local Similarity 43.9%;  Pred. No. 2e-90;
Matches  271;  Conservative  121;  Mismatches  169;  Indels  57;  Gaps  16;
```

```
Qy    73 DNQEDPQTECQQCQRRCRCQQESGPRQQQYCQRRCKEICEE----EEEYNRQ-------- 119
          :| ||| |: ||| |||||||| |||||||||||||||||||    ||||||
Db     9 ENPRDPREQYRQCQEYCRROGQGOGRQQQOQCQQIRCEERLEDQRSQEERERORGRDRQ 68

Qy   120 -RDPQQQYEQCQKHCQRRETEPRHMQTCQQRCERRYEKEKRKQQKRYEEQQQREDEEKYEE 178
           ||||:|||||:|| ||:||||:|||||||||||:||:|:||||||||::||||||:::|
Db    69 PRDPEQRYEQCQQQCE-RQRRGQEQTLCRRRCEQRQOQEERQRQRDRQ---------- 117

Qy   179 RMKEEDNKRDPQQOREYEDCRRRC--EQQEPRQQHQCQLRC-REQQROHGR-GGDMMNPQR 234
          |:::  |||:|||||:||:|||:|  ||||:||:||:||:: ||||:|||| ||::|||
Db   118 --------DPQQ-QYHRCQRRCQIQEQSPEROROQRQQEREQRQYKEQQQGRERGPEASPRR 167

Qy   235 GGSGRYEEGGEEQSDNPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLL 294
          |||||||||||||||||||||:||:|||||:||||||||||||||:|||||||||||||
Db   168 ESRGR---EEQQRHNPYYFHSQSIRSRHESEEGEVKYLERRTERTELLRGTENYRVIL 224

Qy   295 EANPNAFVLPTHLDADAILLVIGGRGALKMIHHDNRESYNLECGDVIRIPAGTTFYLINR 354
          |||||||||||||||||||||||||||||||||:||||||||||||||||||||||||:
Db   225 DANPNTSMLPHHKDAESVAVVTRGRATLTLVSQETRESFNLECGDVIRVPAGATVVVINQ 284

Qy   355 DNNERLHIAKFLQTISTPGQYKEFFPAGGQNP-EPYLSTFSKEILEAALNTQTEKLRGVF 413
          ||||||||:|||||| |||:||||||||||:| ||:||||||||  :|::||||||||||
Db   285 DSNERLEMVKLLQPVNNPGGQPREYYAAGAKSPDQSYLRVESNDILVAALNTPRDRLERFF 344

Qy   414 GQ--OREGVIIRASQEQIRELTRD--DSESRHWHIRRGGESSRGPYNLFNKRPLYSNKYG 469
          |:  |||||||||||:||:||:||  |:||||:||||||||||||||||:|||||||||
Db   345 DQQEOQOREGVIIRASQEKLRALSQHAMSAGQORPW-----GRRSSGPISLKSESPSYSNQFG 400

Qy   470 QAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVVASGEADVEMACPHLS 529
          |||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||||:
Db   401 QFFEACPEEHRQLQEMDLVLNYAEIKRGAMMVPHINSKATVVVVVEGTGRYEMACPHVS 460

Qy   530 GR-HGGRGGGGKRHEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVSSGNENLLFAFGIN 588
          || |||||||||||||||||||||||||||||||:|||||||||||||||||||||||||
Db   461 SQSYEQGGRRROQEEEESTGRPOKVTARLARGDIFVIPAGHPIAITASQNENLRLLGFDIN 520

Qy   589 AQNNHENFLAGRERNVLQQIEPQAMELAFAAAPRKEYEESNSQDQSIFFPGPROHQQQSP 648
          :||||||||||||||||||:|||||||:|||||:|::|||:|:||||||||||||||||
Db   521 GENNQRDFLAG-QNNINQLEREAKELSFNMPREEIEEIFESQMESYFVPPER----QSR 575

Qy   649 RSTKQQQPLVSILDFVGF 666
          :|||||||||||||||:|
Db   576 RGQGRDHPLASILDFAFF 593
```

RESULT 5

```
ID   Q92WI3
AC   Q92WI3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
PV100.
OS   Cucurbita maxima (Pumpkin) (Winter squash).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Cucurbitales; Cucurbitaceae; Cucurbita.
OX   NCBI_TaxID=3661;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KUROKAWA AMAKURI NANKIN; TISSUE=COTYLEDON;
RX   MEDLINE=99107919; PubMed=9891029;
RA   Yamada K., Shimada T., Kondo M., Nishimura M., Hara-Nishimura I.;
RT   "Multiple functional proteins are produced by cleaving Asn-Gln bonds
```

```
PRELIMINARY;    PRT;    810 AA.
```

Query Match

Best Local Similarity

PRELIMINARY;   PRT;   593 AA.

RT   of a single precursor by vacuolar processing enzyme.";
RL   J. Biol. Chem. 274:2563-2570(1999).
DR   EMBL; AB019195; BAA34056.1; -.
DR   HSSP; P02853; 2PHL; .
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PRODOM; PD081059; -; 1.
SQ   SEQUENCE 810 AA; 97314 MW; A829A3F7542266AB CRC64;

Query Match   33.8%; Score 1198.5; DB 10; Length 810;
Best Local Similarity   34.5%; Pred. No. 1.5e-81;
Matches 286; Conservative 141; Mismatches 219; Indels 183; Gaps 23;

```
Qy    1 MAINTSNLCSLLFLLSLFLLSTTVSLAE-------------------------SEF- 31
          ::||  | |                                           ||
Db    1 MALSKVKLRLCLLAFTLFLACLSVGLGDKGESLSSGAGVDHDGCVNRCEELKGKNVDEFA 60

Qy   32 -------------DRQEYEECKRQCMQLETS-GQMRRCVSQCDKRFEE-------------- 65
                      :||   |||:  |      |:| |  | ::|
Db   61 ACKKGGCGVNQRGSPRAEYEVCRLRCQVAERGVEQORKCEQVCEERLREREQGRGEDVDEV 120

Qy   66 --DIDWSKY------------------------DNQEDPQTECQOQCCORRCRQQE----- 93
           :||                            ||:||  :  |  || :|| ||
Db  121 ERRDPEWEREEQORRREHEREREERERRERERERGRGRRDENERDPKREQEEROR--REQE 178

Qy   94 SGPRQQOYCQORCKEICEEEEEYNRQRDPQOQYEQCQOKHCQRRETEPRHMQTCQQRCERR 153
          | |  | || |   |     |||:|||  || || |:  |:|  ||
Db  179 RRRREQEORERERGERDEEDDEN-QRDPDWRREQ----ERREQERRRREQEQEERRERQ 232

Qy  154 Y-------EKE--------------KRKQQKRYEQOQREDEEKYEERMKEEDNKRDP----- 195
          :        |:              :|||  |||  ||:||
Db  233 RRGGRDDEDENQRDPDWRREQKRREQEOERRERERGRGRDDEDENQRDPDWRREQERREQ 292

Qy  196 DCRRRCEQOE---------------PRQQHQCOLRCREQOROHGHRGG-- 227
          | ||| |:|                |:|| |  || |||:|
Db  293 ERRRREREQOERREREHRGGRDVEDENQRDPDWRREQEOERRERERGRGRE 352

Qy  228 -------DMMNPQRGGSGRYEE---------GEEEQSDNPYYFDERSLSTRFRTEEGHISVL 273
                 |  ||||||||||            |||:||||||| :|:|:|||||||| ||
Db  353 EQORSREDDERRRHERQHGGSRVNQVAIRRTEQOSNNPYYFQORFQOSRYRSDEGHRVL 412

Qy  274 ENFYGRSKLLRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHHDNR--- 330
          |:|  ||  |: ||::||| |||:| |: |:||  |  :|:|: ||| | :|  |
Db  413 ERFSERSELLKGIKQNORLALLEARPHFFIVPHHLDAECVLIVVRGRATITTVVQEKRETR 472

Qy  331 -ESYNLECGDVIRIPAGTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPY 389
           ||||  ||| ||||||||:||:|:|  ||| :  :|::|||||:|:||| |||: |:
Db  473 KESYNVESGDVWTIPAGTTLYLANQE-NEDLQIVKLVQPVVNNPGEFKDYLSAGGES-QAY 530

Qy  390 LSTFSKEILEAALNTQTEKLRGVFGQORE--GVIIRASOEQIRELIRDDSESRHWHIRRG 447
           | :|::|||||| || |:| :|  |:|  ||:||||| ||:| || |||||| :||:|
Db  531 YSVFSNDVLEAAALNIPRDKLERIFKQRRERGKIVRASQEOLRALSQ----RATSVKG 585

Qy  448 GESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTR 507
          | || ||:|: ||| : |:|||| :|  |||:|  |::|:||  | :|||||:|:||:|
Db  586 SRGVRAPIKLESQTPVYNNQYGQMFEACPDFEPQLRRTDVATSVVDIKQGGMMVPHFNSR 645

Qy  508 STKVVVVASGEADVEMACPHLSGRHGGRGGGK----RHEEEED-----VHYEQVRARLS 557
          :| |:|  || | |||||||:  :  :|              |:     ||:|||||||
Db  646 ATWVFVSEGAGSFEMACPHIQSSQWQRGRREEEHRWREEEEEERSGRFERVAGRLS 705

Qy  558 KREAIVVLAGHPVFVSSGNENLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAF 617
          :   :|:: ||| : : :|||| :||||||:|||:|||| :| :: ::  || :|||
Db  706 EGGVLVIPAGHPIAIAMASPNENLRLVGFGINAENHHNFLAQRE-NIMNELDREAKELAF 764

Qy  618 AAPRKEVEESNSQDQSIFFPGPRQHQOQSPRSTKQOQPLVSILDFVGF 666
          :  :|||||::||:|| |  || :   |||:|: |||| ||| :||
Db  765 NVEGKQOADEIFRSQRESFFTEGP----EGGRRRSTERSPLLSILKLAGY 809
```

ID   Q43358   PRELIMINARY;   PRT; 525 AA.
AC   Q43358;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN PRECURSOR.
GN   CSV.
OS   Theobroma cacao (Cacao).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Malvales; Malvaceae; Theobroma.
OX   NCBI_TaxID=3641;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAVES;
RX   MEDLINE=92288309; PubMed=1600151;
RA   McHenry L., Fritz P.J.;
RT   "Comparison of the structure and nucleotide sequences of vicilin genes
RT   of cocoa and cotton raise questions about vicilin evolution.";
RL   Plant Mol. Biol. 18:1173-1176(1992).
DR   EMBL; X62625; CAA44493.1; -.
DR   EMBL; X62626; CAA44494.1; -.
DR   HSSP; P02853; 2PHL.
DR   MENDEL; 30919; Thecc;1188;30919.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PRODOM; PD081059; -; 1.
KW   Signal.
FT   SIGNAL   1   24   POTENTIAL.
FT   CHAIN   25   525   VICILIN.
SQ   SEQUENCE 525 AA; 60798 MW; 19114CD5C248905D CRC64;

Query Match   28.9%; Score 1025; DB 10; Length 525;
Best Local Similarity   40.9%; Pred. No. 8.4e-69;
Matches 210; Conservative 93; Mismatches 157; Indels 54; Gaps 11;

```
Qy  109 ICEEEEEYNR---QRDPQOQYEQCQOKHCQRRETEPRHMQTCQQRCERRYEKERRKQQKRY 165
          :|   :|  :  ::|||:||||||| ||  | :| :|| :  |:|| |    |
Db   22 LCSGVSAYGRKQYERDPRQQYEQCQORCCSESEATEEREQOCEQRCER-----EY 70

Qy  166 EEQOREDEEKYEERMKEEDNKRDPQOREYEDCRRRCEQOE--PRQOHQCOLRCREQOROH 223
          :|||  ||:  ||  | ||    |:|  | || | |:|   |:| | |  || ||| |
Db   71 KEQOQOEEEL------------QROYCQOCOGRCQEQOQROQOORCKCWEQOYKEQ 116

Qy  224 GRGGDMMNPQRGGSGRYEEGEEEQSDNPYYFDE-RSLSTRFRTEEGHISVLENFYGRSKL 282
           | |  ::||| || |: :|:::|:||||||    |: |||| ||| || || | ::
Db  117 ER-GEHENYHNHIKKNRSEEEGQORNNPYYFPRRSFSFQTFRFRDEEGNRKILQRPAENSPP 175

Qy  283 LRALKNYRLVLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHHDNRESYNLECGDVIR 342
          |:  :|:||| ||||| |:||:|  || || | : | | |:| :|||||||  : |||
Db  176 LKGINDYRLAMFEANPNTFILPHHCDAEAIYFVTNGKGTIITFVVTHENKESYNVQRGTVVS 235

Qy  343 IPAGTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAAL 402
          :|||:| |:::|:|:|:| |  :|:: : ||:|||||| |:| |||  :|| ||:|  :
Db  236 VPAGSTVVYVSQDNQEKLITIAVLALPVNSPGKYELFFPAGNNKPESYIGAFSYEVLETVE 295

Qy  403 NTQTEKLRGVFGQOR-------EGVIIRASOEQIRELTRDDSESRHWHIRRGGESSRGP 454
          ||| ||| :| :|:|       ||:||| | |:| :    |:   | ||   ||: |
Db  296 NTQREKLEEILEEQRGQKRQOGQOGMFRKAKPEQOIRAISQOATSPRH----RGGE--RLA 349

Qy  455 YNLFNKRPLYSNKYGQAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVV 514
          :|| : ||:|:|: |  :|  ||| :| |:|:| | ||||| || :::   |:|: |:|
Db  350 INLLSQSPVYSNQNGRFFEACPEDFSOFQNMDVAVSAFKLNQGAIFVLPHYNSKATFVFV 409

Qy  515 ASGEADVEMACPHLSGRHGGRGGGK--RHEEEDV-----HYEQVRARLSKREAIVVLA 566
          : |  :|:||||||| | :  :| :  ||:|:|      ||:||:| |||   :|:|
Db  410 TDGYGYAQOMACPHLSRQSOGSOQSOSRQDRREQOEESEEETFGEFQOVKAPLSPGDVFVAPA 469

Qy  567 GHPVFVSSGNENLLFAFGINAQONNHENFLAGR 600
          || | |:| |||| ||||:||:| |:|:||||
Db  470 GHAVTFFASKDQPLNAVAFGLNAQONRIFLAGR 503
```

RESULT 7

ID   Q03865   PRELIMINARY;   PRT;   582 AA.
AC   Q03865;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN-LIKE EMBRYO STORAGE PROTEIN.
GN   GLB1-L.
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=W64A;
RX   MEDLINE=92090707; PubMed=1752424;
RA   Belanger F.C., Kriz A.L.;
RT   "Molecular basis for allelic polymorphism of the maize Globulin-1
RT   gene.";
RL   Genetics 129:863-872(1991).
DR   EMBL; X59083; CAA41809.1; -.
DR   HSSP; P50477; 1CAU.
DR   MENDEL; 11234; Zeama;1188;11234.
DR   INTERPRO; IPR000901; -.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
DR   PRODOM; PD081059; -; 1.
KW   Seed storage protein.
SQ   SEQUENCE 582 AA; 66162 MW; 680D85FECD7CB885 CRC64;

Query Match 25.4%; Score 898; DB 10; Length 582;
Best Local Similarity 39.2%; Pred. No. 3.2e-59;
Matches 197; Conservative 91; Mismatches 184; Indels 30; Gaps 9;

Qy   182 EEDNKRDPQQREYEDCRRRCEQQEPRQQHQCQLRCREQQRQHGRGGDMMN-PQRGGSG-- 238
         :::::   :::   :::::::::::::::::::::::::::::::::::: ::::::::
Db   25 EDDNHHHGHGKSGQCVRRCEDRRWHQRPRCLEQQCREEEREKRQERSRHEADRSGEGSS 84

Qy   239 ---RYEEGEEEQSDNPYYFDERSLSTRFRTEEGHIISVLENFYGRSKLLRALKNYRLVLLE 295
            :::: :::::::: ::::::::::::: :::::::::::::::::::::::::::::
Db   85 EDEREQEKEKQKDRRPYVFDRRSFRRVVRSEQGSLRVLRPFDEVSRLLRGIRDYRVAVLE 144

Qy   296 ANPNAFVLPTHLDADAILLVIGRGALKMIHHDNRESYNLECGDVIRIPAGTTFYLINRD 355
         :: ::::
Db   145 ANPRSFVVPSHTDAHCICYVAEGEGVVTTIENGERRSYTIKQGHVFVAPAGAVTYLANTD 204

Qy   356 NNERLHIAKFLQTISTPCQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLRGVFGQ 415
         ::::::::::::::: : ::::::::::::::: ::::::::::::::::::::::: ::
Db   205 GRKKLVITKILHTISVPGFQFFPGPGCRNPESFLSSFSKSIQRAAYKTSSDRLERLFGR 264

Qy   416 --QREGVIIRASQEQIRELTRDDSE--- -SRHWHIRRGGESSRGPYNLFNKRPLYSNKYGQ 470
           ::::::::::::::::::::::::   ::::::::::::::::::::::::::::::::
Db   265 HCQDKGIIVRATEEQTRELRRHASEGGHGPHWPLPPGE-SRGPYSLLDQRPSIANOHGQ 323

Qy   471 AYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSG 530
         ::
Db   324 LYEADARSFHDLAEHDVSVSPANITAGSMSAPLFNTRSFKIAYVPNGKGYAEIVCPHRQS 383

Qy   531 RHG----GRGGGKRHEEEEDVH-------------------YEQVRARLSKREAIVVLAGHPVVFVSSG 576
         :::   ::::::::::::: :   ::::::::::::::::::::::::::::
Db   384 QGGESERERDKGRRSEEEEESSEEEEQAGQGYHTIRARLSPGTAEVVPAGHFFVAVASR 443

Qy   577 NENLLFAFGINAQONNHENFLAGRERNVLQQIEPQQAMELAFAAPRKEVEEESFNSQDQSIF 636
         :: : ::   ::
Db   444 ,DSNLQIVCFEVHADRNEKVFLAGAD-NVLQKLDRVAKALSFASKAEEVDEVLGSRREKGF 502

Qy   637 FPGPRQ---HQQQSPRSTKQQQ 655
         :::   ::::
Db   503 LPGPEESGGHEBREQEEEEREE 524

RESULT 8

ID   Q03678   PRELIMINARY;   PRT;   637 AA.
AC   Q03678;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   FETAL STORAGE PROTEIN.
GN   BEG1 OR GLB1.
OS   Hordeum vulgare (Barley), and Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
OX   NCBI_TaxID=4513, 4565;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93287988; PubMed=8510647;
RA   Heck G.R., Chamberlain A.C., Ho T.H.D.;
RT   "Barley embryo globulin 1 gene, Beg1: characterization of cDNA,
RT   chromosome mapping and regulation of expression.";
RL   Mol. Gen. Genet. 239:209-218(1993).
DR   EMBL; M64372; AAA32936.1; -.
DR   EMBL; M81719; AAA34269.1; -.
DR   HSSP; P02853; 2PHL.
DR   MENDEL; 8553; Horvu.1188;8553.
DR   INTERPRO; IPR000901; -.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
DR   PRODOM; PD081059; -; 1.
KW   Seed storage protein.
SQ   SEQUENCE 637 AA; 72252 MW; F323F4FF99947C3C CRC64;

Query Match 24.4%; Score 864.5; DB 10; Length 637;
Best Local Similarity 35.1%; Pred. No. 1.2e-56;
Matches 222; Conservative 98; Mismatches 219; Indels 93; Gaps 20;

Qy   70 SKYDNQEDPQ--TECQQCQRRCRQOESGPROQQVCQRRCKEICEEEENRQRDPQQYE 127
         :::   : ::::::::::::::::::::::::::: :::   ::::: :::::::
Db   27 ASHDDEDDRRGGHSLQOQCVQRCRQER--PR---YSHARCVQEC------- RDDQOQH- 71

Qy   128 QCQKHCQRRETEPRIMQTCQQRCEBRYEKEKRKQQKRYEQQOREDEEKYEERMKEEDNKR 187
         ::
Db   72 ------------GRHEQEEQGRGRGWHGEGEREEEHGRGRGRHCEGEREEEHGRGRH 119

Qy   188 DPQQREYEDCRRRCEQQOEPRQOHQCQLRCREQQROHGRG----GDMMNPQRGGSGRYEEG 243
         ::::::::::::::::::::::::::::::::::::::   ::::::::::::::::::
Db   120 GEGEBEEEERGRGHGRHGEGE---------REEERGRGRHGEGEREEEEGRGRGRGRGEG 170

Qy   244 E---EEQSDN--PYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRALKNYRLVLLEANPN 299
         :   :::: :::::::::::::: :::::::::::::::::::::::::::::::::::
Db   171 ERDEEOGDSRRPYVFGRPRSFRRIIQSDHGFVRALRPFDQVSRLLRGIRDYRVAIMEVNPR 230

Qy   300 AFVLPTHLDADAILLVIGRGALKMIHHDNRESYNLECGDVIRIPAGTTFYLINRDNNER 359
         :: ::::::
Db   231 AFVVPGFTDADGVGVVAQGEGVLVIENGEKRSYTVKEGDVIVAPAGSIMHLANTDGRRK 290

Qy   360 LHIAKFLQTISTPCQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLRGVFGQQ-- 416
         :::::::::::::: : :::::::::::::::: :::::::::::::::::: ::
Db   291 LVIAKILHTISVPGKF-QFL----SVKPLLASLSKRVLRAAFKTSDERLERLENQRGQ 344

Qy   417 ---REGVIIRASOEQIRELTRDDSE----SRHWHIRRGGESSRGPYNLFNKRPLYSNKYGQ 470
            ::::::::::::::::::::::   ::::::::::::::::::::::::::::::::
Db   345 EKIRSVSIVRASEEQLRELREAAEGGQGHRWPLPPFRGDSRDTFNLLEQRPKIANRHGR 404

Qy   471 AYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSG 530
         :: :::::::::::::::::::
Db   405 LYEADARSFHALANQDVRVAVANITPGSMTAPYLNTQSFKLAVVLEGEGVQIVCPHL-G 463

Qy   531 RHG---GRGGGKRHEEEED------------------VHYEQVRARLSKREAIVVLA 566
         :::   ::::::::::::   :: ::::::: :::::::
Db   503 LPGPEESGGHEBREQEEEEREE 524

```
Db   464 RESESEREHGKGRRREEEEDDQRQQRRGSESESESEEEEEQQQRYETVRARVSRGSAFVVPP 523
Qy   567 GHPVVFVSS--GNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAAPRKEV 624
         :|:|:|||   |||| ||||||||||| ||| ||||||  |||||| |||||||||||
Db   524 GHPVVELSSSQGSSNLQVVCFEINAERNERVWLAGR-NNVIGLGSPAQELTFGRPAREV 582
Qy   625 EESFNSQODQSI-FPFGPRQHQOQSPRSTKQQQ 655
         |:|||||| |:| :|||||||| ||| :||
Db   583 QEVFRAQDQDEGFVAGP---EQQSREQEQEQE 611
```

RESULT 9

```
ID   Q9M3X6
AC   Q9M3X6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   CONVICILIN PRECURSOR.
GN   CVC.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. BIRTE; TISSUE=SEED COTYLEDON;
RA   Casey R.;
RT   "The sequence of a pea convicilin cDNA.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ276875; CAB82855.1; -.
KW   Signal.
FT   SIGNAL        1     29       POTENTIAL.
FT   CHAIN        30    613       CONVICILIN.
SQ   SEQUENCE   613 AA;  72063 MW;  49AED99F6135DD19 CRC64;
```

Query Match Best Local Similarity 30.5%;
Matches 212; Conservative 130; Mismatches 232; Indels 122; Gaps 15;

Score 842; DB 10; Length 613; Pred. No. 5.4e-55;

```
Qy     1 MAINTSNICSLLFLLSLLSTTVSLAESEFDRQEYECKRCQMQLETSGQMRRCVSQCD 60
         |:  : :|: |  :|::|||||: |
Db     1 MATTIKSRFPLLLLLGIIFLASVV------------------------CVT--- 27

Qy    61 KRFEEDIDWSKYDNQEDPQTECQCOCQRCRQESGPRQQQYCQRRCKEICEEEEEYNRQR 120
         :: ::  | |: |||||:  ||| |:|:|| | : ||:
Db    28 ----YANVDEGSEPRVPAQRERGRQEGEKEEKRHGEWRPSYEKE--EDEEEGQRER 77

Qy   121 DPQQQYEQCQHCQRRETEPRHMQTCCQRCERRYEKRKQQKRYEEQQQREDEEKYEERM 180
         : | |:::: : | |:  ::   ||: :|||
Db    78 GRQEGEKEKRIGEWRPS------------YEKQEDEEEKQKYRQREKEDE-EEKQ 121

Qy   181 KEEDNKRDPQQREYEDCRRCEQOQEPRQQ----------------HQCQLRCEQQRQHGR 225
         ::|:  ||  | :| |||::||     ::|       ||:|  |||:  ||
Db   122 KYQYQREKKEQKEVQPQRERWEREEDEEQVDEEWRGSQRREDPEERARLRHREERTKRDR 181

Qy   226 GGDMMNPQRGGSGRYEEBGEEEQSDNPYYFDERSLSTFRFRTEEGHISVLENFYGRSKLLRA 285
         :: |  :  |  | ::: ::|:: |||  ||||||:|||| || : :| :|
Db   182 ----RHQRGEGEEERSSESQERRNPFLFKSNKFLTLFENENGHIRLLQRFDKRSDLFEN 236

Qy   286 LKNYRLVLLEANPNAFVLPTHLDADALLLVIGGRGALKMIHIDNRESYNLECGDVIRIPA 345
         | ||||: ::  |: :|||:::|||:|||| ::| |:|  ||||:| |||:|::|::||
Db   237 LQNYRLVEYRAKPHTIFLPQHIDADLLLVVLSGKAILTVLSPNDRNSTNLERGDTIKLPA 296

Qy   346 GTTFYLINRDHNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQ 405
         |||:||:|:| :|:|:::|:||||:|:|||||||:|||||||||||||:||:||||||||
Db   297 GTTSYLVNQDDEEDLRLVDLVIPVNGPGKFEAFDLA--KNKNQYLRGFSKNILEASYNTR 354

Qy   406 TEKLRGVFGQQRE------------------GVIIRASQEQIRELTRDDSESRHWHIRGGESSRG 453
         |:::|:||  :|        ||:||||:||||:|||||:|| : |
Db   355 YETIEKVLLEEQEKDRKRRQGEETDAIVKVSREQIELKLLAKSSS---KKSLPSEFE 410
```

RESULT 10

```
ID   022120
AC   022120;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ALPHA SUBUNIT OF BETA CONGLYCININ.
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. WASESUZUNARI;
RA   Maruyama N., Katsube T., Wada Y., Utsumi S.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB008678; BAA23360.1; -.
DR   MENDEL; 25074; Glyma;1188;25074.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00190; Seedstore_7s; 1.
DR   PRODOM; PD081059; -; 1.
DR   HSSP; P50477; 1CAU.
SQ   SEQUENCE   544 AA;  63296 MW;  04B243929A9C9AB CRC64;
```

Query Match Best Local Similarity 32.4%;
Matches 188; Conservative 126; Mismatches 197; Indels 70; Gaps 13;

Score 812; DB 10; Length 544; Pred. No. 8.2e-53;

```
Qy   107 KEICEEEEEYNRQRDPQQQYEQCQCHCQRRETE------------PRHMQTCCQRCERRYEKE 157
         ::|| ::: | : :  |:: ||| |:|| |:||        ||:||| |||:|||:
Db     4 KEECEEGEIPRPRPQHPEREPQOPGEKEDEDEOPRPIFPRPQOPEEEHEQREEQE 63

Qy   158 -KRKQQKRYEEQQQREDEEKRYEERMKEEDNKRDPQQREYEDCRRCEQOQEPRQQHQCQLRC 216
         :|::  ||:|   :|::  :::: |:||:  :|| ||| :|:| |||:|||   :|||:|
Db    64 WPRKEEKRGEKGSEEEDE------DEDEOEDERQFPFP---RPHQKEERKQEEDE--D 111

Qy   217 REQQRQHGRGGGDMMNPQRGGSGRYEEBGEEEEQSDNPYYFDERSLSTFRFRTEEGHISVLENF 276
         :|:|:|:  ::  |  :  |  | ::: :::|: |||  ||||||:||||| || : :| :|
Db   112 EEQQRESEESED------------SELRRHKNKNPFLFGGSNRFETLFKNQYGRIRVLQRF 159

Qy   277 YGRSKLLRALKNYRLVLLEANPNAFVLPTHLDADALLLVIGGRGALKMIHIDNRESYNLE 336
         ||||| | |:|||||: ::  |: :|||:::|||:|||| ::| |:|  ||||:| |||
Db   160 NQRSPQLQNLRDYRILEFNSKPNTLLLPNHADADYLIVILNGTAILSLVNNDDRDSYRLQ 219

Qy   337 CGDVIRIPAGTTFYLINRDHNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKE 396
         :|:|::||||||:||:|:| :|:|:::|:||||:|:|||||||:|||||||||||||:||
Db   220 SGDALRVPSGTTYYVVNPDNNENLRLITLAIPVNKPGRESFFLSSTEAQOSYLQGFSRN 279

Qy   397 ILEAALNTQTEKLRGVF------------------GQOR--EGVIIRASQEQIRELTRDDSESRHWHIR 445
         ||||:|: |::::|:||        ||:|  ||:||||||:|| : |
Db   280 ILEASYDIKFEEINKVLFSREEGQOOGEORLQESVIVEISKEQIRALSKRAKSSS----R 335
```

```
Qy  446 RGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFN 505
            ||| |    ||||  | ||||: | |::  |:: |   ||: ||:|:||:|||:||:|| ||
Db  336 KTISSEDKPFNLRSRDPIYSNKLGKFFEITPEKNPQLRDLDIFLSIVDMNEGALLLPHFN 395

Qy  506 TRSTKVVVVASGEADVEMACPHLSGRHGGRGGGK------RHEEEEDVHYEQVRARLSKRE 560
                                  :  |||:|| ::|     : |::| ::   |  ||::
Db  396 SKAIVLVINEGDANIELV--------------GLKEQOQEQOOEOPLEVRKYRAELSEQD 443

Qy  561 AIVVLAGHPVVFVSSGGNENLLLFAFGINAQNHHENFLAGRERNVLQQIEPQAMELAFAAP 620
            ||:: ||||||  |:| ||| ||:| ||||| | : |||| ::|| |  |||:||||| | 
Db  444 IFVIPAGYPVVV--NATSNLNFFAIGINAENNQRNFLAGSQDNVISQITPSQVQELAFPGS 501

Qy  621 RKEVEESFNSQDQSIFFPGPRQHQQOSPRSTKQQOQPLVSIL 661
            :|:|:|: ||| | |||||:: : | |||| | :: |: ||
Db  502 AQAVEKLLKNQRESYFVDAQPKKEEGNKGRK--GPLSSIL 540
```

RESULT 11

```
ID   Q9SP11          PRELIMINARY;      PRT;   489 AA.
AC   Q9SP11;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   SUCROSE BINDING PROTEIN HOMOLOG S-64.
GN   SBP.
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pedra J.H.F., Delu-Filho N., Pirovani C.P., Contim L.S., Dewey R.E.,
RA   Otoni W.C., Fontes E.P.B.;
RT   "Antisense and sense expression of a sucrose binding protein homologue
RT   gene from soybean in transgenic tobacco affects plant growth and
RT   carbohydrate partitioning in leaves.";
RL   Plant Sci. 0:0-0(1999).
DR   EMBL; AF191299; AAF05723.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
SQ   SEQUENCE   489 AA;  55834 MW;  9BBC0D45EDECECD2 CRC64;
```

```
Query Match          22.7%;  Score  802.5;  DB 10;  Length 489;
Best Local Similarity 35.8%;
Matches  175;  Conservative  98;  Pred. No. 3.6e-52; Mismatches 171;  Indels  45;  Gaps  11;
```

```
Qy  179 RMKEEDNKRDP------------QQREYED-------CRRRCEQOEPRQOHQCOLRCREQQRQH 223
            :||:     |            | |::|:       |::|   |  :   |   | :|||
Db   27 KLKETEVEEDPELVTCKHQCOQOROYTESDKRTCLQQCDSMKQEREKQVEEETREKEEEH 86

Qy  224 GRGGDMMNPQRGGSGRYEEGEEEEQSDNPYVFDE-RSLSTRFRTEEGHIISVLENFYGRSKL 282
            |:      :| |  ||  | |:| ::||||| || ||  |||:  | :| ||:||||||||
Db   87 -------OEQHEEEQDQNPYVFEEDKDFSTRVETEGGSIRVLKKFTEKSKL 130

Qy  283 LRALKNYRLVLLLEANPNAFVLPTHLDADAILLVIGGRGALKMHHDNRESYNLECGDVIR 342
            | ::||:||: ||||  :||||| |:||: ||||| |: ::  | | |:| :| |||:|
Db  131 LQGNENFRLAILEARAHTFVSPRHFDSEVVLTNIKGRAVLGLVRESETEKITLEPGDMIH 190

Qy  343 IPAGTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAAL 402
            |||| ||:|:||:|||:| || ||:| ||||: ||||  || ||| |:|||| :||:||
Db  191 IPAGTPLYIVNRDENEKLLLAMLHIPVSTPGKFEEFFGPGGRDPESVLSAFSWNVLQAAL 250

Qy  403 NTQTEKLRGVFGQQREQVIIRASQEQIRELTRDDSESRHWHIRRGGESSSRGPYNLFNKRP 462
            :||  |||: : || :  :: |: | | |  |  ||  ::   :|| |  |  ||:|:||
Db  251 QTPKGKLERLENQQNEGSIFKISRERVRALA-PTKKSSWPF--GGE-SKAQFNIFSKRP 306

Qy  463 LYSNKYGQAYEVKPEDYRQ-LQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADV 521
            ::|| ||: :||:|:| :   |::: | | ||:||:||   | |   |  |: :  |  :|
Db  307 TFSNGYGRLTEVGPDDEKSWLQRLNLMLTFTNITQRSMSTIIHYNSHATKIALVMDGRGHL 366
```

RESULT 12

```
Qy  522 EMACPHLSGRHGGRGGGKRHEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVSSGNENLL 581
            :: |||:| | : ||::|   :| || || ::   | |:: | || ||:|: : |::|||
Db  367 QISCPHMSSRSDSK-----------HDKSSPSYHRISADILKPGMVFVVPGCHPFVTIASNKENLL 420

Qy  582 LFAFGINAQNHHENFLAGRERNVLQQIEPQAMELAFAAPRKEVEESFNSQDQSIFFPGPR 641
            :   :|  |::  |||| | ||::  : |||:||| ||  |::| :| :    ||||  |
Db  421 IICFEVNVRDNKKFFFAGKD-NIVSSLDNVAKELAFNYPSEMVNGVF-ERKESLFFPEEL 478

Qy  642 QHQQOSPRS 650
            | :  : |
Db  479 PSEERGRRA 487
```

```
ID   Q41674          PRELIMINARY;      PRT;   545 AA.
AC   Q41674;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   CONVICILIN PRECURSOR.
OS   Vicia narbonensis.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Vicia.
OX   NCBI_TaxID=3912;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=COTYLEDON;
RA   Koch G., Koenig S., Becker C., Horstmann C., Schlesier B.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z71986; CAA96513.1; -.
DR   HSSP; P50477; 1CAU.
DR   MENDEL; 12432; Vicna:1188;12432.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PRODOM; PD081059; -; 1.
KW   Signal; Seed storage protein.
FT   SIGNAL        1     28       POTENTIAL.
FT   CHAIN        29    545       7S GLOBULIN, CONVICILIN.
SQ   SEQUENCE   545 AA;  62810 MW;  459A876F92F5A87E CRC64;
```

```
Query Match          22.2%;  Score  787;  DB 10;  Length 545;
Best Local Similarity 33.5%;
Matches  191;  Conservative 102;  Pred. No. 6.1e-51; Mismatches 208;  Indels  70;  Gaps  14;
```

```
Qy  109 ICEEEEEYNRQRDPQQQYEQCQKHCQRRETEPRIHMQTCQQRCERRYEKERKQQKRYEE- 167
            :| :    :| ||| |||:  |:||| |||  :|::|:|| |||||:|:|::||||:||
Db   23 LCVTYANYDEGTEPRVPGQRERGRQEGEKEEKRH-----------GEWRPSHEKEAQPGRRERW 75

Qy  168 QQREDEEKYEERMKEEDNKRDPQQREYEDCRRRCEQQEPRQQHOCOLRCREQQROHGRGG 227
            :  |:||  |   |:|   || : :  |: |  |||: |  |||  |  | :|||:||||
Db   76 ETSEEEERVDEEWRGSQRHEDPEERARE--RYRAEEERRRQWE---------------- 117

Qy  228 DMMNPQRGGSGRYEEGEEEQSD--NPYYFDERSLSTRFRTEEGHISVLENFYGRSKLLRA 285
            :  |   |:  :|| | |:|    ||:: |:|| || |: |: |||  |:|| | |:| 
Db  118 -------------GEEKEGSSKSQERRNPFLFKSNKFLTLFENENGHIRRLQRFDKRSDLFEN 167

Qy  286 LKNYRLVLLLEANPNAFVLPTHLDADAILLVIGGRGALKMHHDNRESYNLECGDVIRIPA 345
            |:|||||   | |:  ||||||:||| |||||: ||| ::  |:|:|||| | || |:||
Db  168 LQNYRLVEYRAKPHTIFLPQHIDADLILTVLSPNDRNSYNLERGDTIKLPA 227

Qy  346 GTTFYLINRDNNERLHIAKFLQTISTPGQYKEFFPAGGQNPEPYLSTFSKEILEAALNTQ 405
            |||:|::|:|: | |: || | ||: |: | |: |: || |||||:|||:|||||:|||:
Db  228 GTTSYLLNQDDDEDLRVVDLSISVNRPGKVESFGLSGSKN--QYLRGFSKNILEASLNTK 285

Qy  406 TEKLRGV---------FGQQRE------GVIIRASQEQIRELTRDDSESRHWHIRRGGESS 451
             |:::::         :||::|      || :|| : ::| ::: : | :     ::| ||
Db  286 YETIEKVLLEEPQOSIGQKRRSQROETNALVKYSREQVEELKLRLAKSSS----KKGVSSE 341

Qy  452 RGPYNLFNKRPLYSNKYGQAYEVKPE-DYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTK 510
Db                                                              510
```

RESULT 13

```
ID   O49927
AC   O49927;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   P54 PROTEIN.
GN   P54.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fabales; Fabaceae; Papilionoideae; Pisum.
     NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Castillo J., Marquez J.A., Franco L., Ballestar E., Rodrigo M.I.;
RL   Submitted (FEB-1997) to the EMBL/Genbank/DDBJ databases.
     EMBL; Y11207; CAA72090.1; -.
     HSSP; P50477; 1CAU.
     MENDEL; 28319; Pissa;3005;28319.
     INTERPRO; IPR000005; -.
     INTERPRO; IPR001113; -.
     PFAM; PF00546; Seedstore_7s; 1.
     PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
     PRODOM; PD081059; -; 1.
SQ   SEQUENCE 483 AA; 54662 MW; 8127BDAAA0178F3D CRC64;

PRELIMINARY;    PRT;    483 AA.

Query Match          21.7%;  Score 769.5;  DB 10;  Length 483;
Best Local Similarity 31.4%;  Pred. No. 1.1e-49;
Matches 162;  Conservative 114;  Mismatches 169;  Indels 71;  Gaps 12;
```

```
Qy  342 FEPFNLRSQNPKYSNKPGKLFEITPERKYYPQLQDLDIFVSSVEINEGGLMLPHYNSRAIV 401
        |::|   | ::| |  :|   |   |:::     |  ||  :|:||: ::  :|| |:|  |
Db  511 VVVVASGEADVEMACPHLSGRHGGRGGGKRHEEEEDVHYEQVRARLSKREAIVVLAGHPV 570

Qy  402 ILLVNEGKGNLE-----LVGLKNEQEQOREREDEQQV--QRYEARLSPGDVVIIPAGHPV 454
        ||  | :|:|:|     | |:|||:|: ::: :|:| |   ::| ||:|::||||||:| |
Db  571 VFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAAPRKEVEESFNS 630

Qy  455 AVSASSNLNLL--GFGINAENNQRNELTGSDDNVISQIENPVKELTFPGSAQEVNRLLKN 512
        |  ||:|||||  ||||||:||||||:|  ||  :|||||||:|| |||| |:|||||:||
Db  631 QDQOSIFFPGPRQHQQOQSPRSTKQQOQPLVSIL 661

Qy  513 QEHSHF--ANAEPEQKGEESQRKRSPISSIL 541
        ||:|||  ||||||:|||||:|||||||:||
Db      QDQOSIFFPGPRQHQQOQSPRSTKQQOQPLVSIL 661
```

```
Qy  135 RRETEPRHMQTCQORCERRYEKEKRKQOKRYEEQOREDEEKYEERMKEEDNKRDPQQOREY 194
        ||:| || | |  |:   |  ::|:||:|| :  : | |:|  :|:  |||:||:  ||||
Db   28 RKEKDP-ELTTCKDQCD----------MORQDEEDK------------------------ 53

Qy  195 EDCRRCEQOEPRQOHQCQLRCEROQROHGRGGDMMNPQRGGSGRYEEGEEEQSDNPYYF 254
        ||   || |:| :|:  | | | :| :|:|||  ||||||:||:|||:|||||||||||
Db   54 RICMERODYIKKQE-----------------RQKHKEH--------------------- 93

Qy  255 DERSLSTRFRTEEGHIISVLENFYGRSKLLRALKNYRLVLLEANPNAFVLPTHLDADAILL 314
        ||||||||||:||||||||||||||||||||||||||:|:||||||||:|:||||||||||
Db   94 EDNDFETKIDTKDGRVLILNKFNEKSKLLKNIENYGLAVLEIKANAFLSPHHYDSEAILF 153

Qy  315 VIGGRGALKMIHHDNRESYNLECGDVIRIPAGTTFYLINRDNNERLHIAKF---LQTIST 371
        |:|||||||||||||||||||||||:||||||||||:|::|:|||||:|||   |||:||
Db  154 NIKGRGIIGLVAEDRTERFNLEEGDIMRVPAGTPMYLVNRDENEKLYIAAFHMPPSGSA 213

Qy  372 PGOYKEFFPAGGQNPEPYLSTFSKELLEAALNTQTEKLRGVFGQOREGVIIRASQEOIRE 431
        |||||||||||||||||||||||||||||||||:|:||||||:|:|||||||:|||:|||
Db  214 PVNLEPFFESAGRKPESVLNTFSSKVLQAALKSSKGELETVLDEOKRGRIFKIEKEDVRG 273

Qy  432 LTRDDSESRHWHIRRGGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDYRQ-LQDMDLSVF 490
        |:||||:|||||||||||:|||:||||||||||||:|||||||||||||:  ||||||||
Db  274 LAPKKS----LWPF---GGPFKSPENIFSNNPAFSNKFGSLFEVGPSQOEKSGLEGLNLMLT 327
```

RESULT 14

```
ID   Q40873
AC   Q40873;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN-LIKE STORAGE PROTEIN.
OS   Picea glauca (White spruce).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Coniferophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
     NCBI_TaxID=3330;
RN   [1]
RP   SEQUENCE FROM N.A.
RT   "Vicilin-like seed storage proteins in the gymnosperm interior spruce
RT   (Picea glauca/engelmanii).";
RA   Newton C.H., Flinn B.S., Sutton B.C.;
RL   Plant Mol. Biol. 20:315-322(1992).
RX   MEDLINE=93004485; PubMed=1391775;
RC   STRAIN=PG118; TISSUE=SOMATIC EMBRYO;
     EMBL; X63191; CAA44873.1; -.
     HSSP; P02853; 2PHL.
     MENDEL; 12115; Picgl;1188;12115.
     INTERPRO; IPR001113; -.
     PFAM; PF00546; Seedstore_7s; 1.
     PRODOM; PD081059; -; 1.
KW   Seed storage protein.
SQ   SEQUENCE 448 AA; 50200 MW; 467AE5BA12E6D157 CRC64;

PRELIMINARY;    PRT;    448 AA.

Query Match          21.2%;  Score 752;  DB 10;  Length 448;
Best Local Similarity 37.8%;  Pred. No. 2e-48;
Matches 162;  Conservative 85;  Mismatches 158;  Indels 24;  Gaps 9;
```

```
Qy  491 IANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKRHEEEEDVHYE 550
        |||:|||||||||||||||||||||||||||||||||||||||||||||||||:  :  |
Db  328 LANITKGSMSTIHYNTNANKIALVIDGEGELEMACPHMPS-----SSSNSRQKKSSISYH 382

Qy  551 QVRARLSKREAIVVLAGHPVVFVSSGNENLLLFAFGINAQNNHENFLAGRERNVLQQIEP 610
        ||||||||||||||||||||:||||||||||||||||||||||||||||||||:| :|:|
Db  383 NINAKLRPGVMFVVPAGHPFVNIASKKNLIVVCFEVNAQRNKLALAGK-KNIVSALDK 441

Qy  611 QAMELAFAAPRKEVEESFNSQODQOSIFFPGPRQHQQO 646
        |||||||||||:||||||||:|||||||||||||||||
Db  442 AAKEVAFDIAAEKVDEVEERKEE-FFFPVDNEERKE 476

Qy  232 PQRGGSGRYEEGEEEQOSDNPYYFDEERSLSTRFRTEEGHIISVLENFYGRSKLLRALKNYRL 291
        |:|:|||||:||||:|:|||||||::||:||:|||||||||||||||||||||||||||| |
Db  153 VNDDHNTLRIASLVRPVSTVRGEYQPFYVAGGRNPQTYVSAFSDDVLEAAFNTNVQOLE 212

Qy  292 VLLEANPNAFVLPTHLDADAILLVIGGRGALKMIHHDNRESYNLECGDVIRIPAGTTFYL 351
        |||||||||:|:|:||||||||:|:|||||||||||||||||:||||||||||||||||||
Db  213 RIFGGHKSGVIIHANEEQIREMKR----------------KRGFSAGSMSAPEHPKPFNLRNQKPD 262

Qy  352 INRDNNERLHIAKFLQTIST-PGQYKEFFPAGGQNPEPYLSTFSKELLEAALNTQTEKLR 410
        ||:|:|||||:|||||:|:| ||||||||||||||||||||||||||||||||:|:|||
Db  263 FENENGRFTIAGPKNYPFLDALDVSVGLADLNPGSMTAPSLNSKSTSIGIVINGEGRIEM 322

Qy  411 GVFGQOREGVIIRASQEQIRELTRDDSESRHWHIRRG---GESSR----GPYNLFNKRPL 463
        ||:|:|||||||||:|||:||| |||||:|||||||||   |:|:|    ||||||||||
Db  323 ACPHLSGRHGGRGGGKRHEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVSSGNENLLLF 383

Qy  464 YSNKYGQAYEVKPEDYRQLQDMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEM 523
        :|||:|||||||||||:|:||||||||:|||:|||||||||||||||||||||||||||||
Db  384 AFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAAPR-KEVEESFNSQODQOSIFFPGPRQ 442

Qy  524 ACPHLSGRHGGRGGGKRHEEEEDVHYEQVRARLSKREAIVVLAGHPVVFVSSGNENLLLF 583
        |||||||:||||||||||||||||||||:|||:||||||||||||||||||||||||:||
Db  443 ACPIL--GQIG--WSSPERGDQOITYQRVWAKLRTGSVIVPAGHPITEIASTNSRLQIL 379

Qy  584 AFGINAQNNHENFLAGRERNVLQQIEPQAMELAFAAPR-KEVEESFNSQODQOSIFFPGPRQ 642
        :|:|   :  ::| ||| |:|:   |  ||  :|||||
Db      ACPIL--GQIG--WSSPERGDQOITYQRVWAKLRTGSVIVPAGHPITEIASTNSRLQIL 379
```

```
Db    380 WFDLNTRGNERQFLAGK-NNVLNTLEREIRQLSFNVPRGEEIEEVLQAQKDQVILRGPQR 438

Qy    643 HQQQSPRST 651
              ||:
Db    439 RSRDEARSS 447


Qy    662 DFVGF 666
              |||
Db    482 DFAGF 486

Search completed: March  1, 2001, 16:08:56
Job time: 1540 sec


RESULT 15
Q9LUJ7
ID   Q9LUJ7;        PRELIMINARY;    PRT;    486 AA.
AC   Q9LUJ7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   GENOMIC DNA, CHROMOSOME 3, P1 CLONE: MWI23.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RA   Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=COLUMBIA;
RA   Nakamura Y.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT   features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT   clones.";
RL   DNA Res. 7:131-135(2000).
DR   EMBL; AB022223; BAB01239.1; -.
SQ   SEQUENCE  486 AA;  55063 MW;  325ECF68D9A6345B CRC64;


Query Match            21.1%;  Score 749;  DB 10;  Length 486;
Best Local Similarity  35.9%;  Pred. No. 3.7e-48;
Matches  174;  Conservative  92;  Mismatches 179;  Indels  40;  Gaps  12;

Qy    198 RRRCEQOEPRQHQCQLRCREQQRQHGRGGDMMNPQR---GGSGRYEEGEEEQSDNPYYF 254
              : ||  ||: ::|| | |                   ::  ||| :||:| | |
Db     26 RVEIDEFEPPQQ-------------------------GEQEGPRRRPGGGSS--EGWEEESTNHPYHF 66

Qy    255 DERSLSTRFRTEEGHISVLENFYGRS-KLLRALKNYRLVLLEANPNAFVLPTHLDADAIL 313
              :||| :|::||| ||  |    :||| |:|| | ||  ||  ||: | |||||||||:
Db     67 RKRSFSDWFQSKEGFVRVLPKFTKHAPALFRGIENYRFSLVEMEPTTFFVPHHLDADAVF 126

Qy    314 LVIGGRGALKMIHHIDNRESYNLECGDVIRIPAGTTFYLINRDNNERLHIAKFLQTISTPG 373
              :||:||| ||:|:: |  |    |||:| |:  ||||  ||| | ||  | | ::|  ||
Db    127 IVLQGKGVIEFVIDKTKESFHITKGDVVRIPSGVTNFITNTNQTVPLRLAQITVPVNPG 186

Qy    374 QYKEFFPAGGQNPEPYLSTFSKEILEAALNTQTEKLRGVFGQQRE---GVIIRASQEQIR 430
              | || : : :|   | |   |  ::|  |   |    |  |  |     |  | : |
Db    187 NYKDYFPAASQFQQSYENGFTKEVLSTSFNVPEELLGRLVTRSKEIGQGIIRRISPDIK 246

Qy    431 ELTRDDSESRHWH-IRRGGESSRG---------PYNLFNKRPLYSNKYGQAYEVKPEDYRQLQ 483
              || :          |  | ||          ::||||  | :|:| :|  | :| :| |||
Db    247 ELAEHATSPSNKHKAKKEKEEDKDLRTLWTPENLFAIDPIYSNDFGHFHEAHPKNYNQLQ 306

Qy    484 DMDLSVFIANVTQGSMMGPFFNTRSTKVVVVASGEADVEMACPHLSGRHGGRGGGKRHEE 543
              | ::|| |  | :||:||: :|||||    ||||:|||  ||| | | |:| |:| ||::
Db    307 DLHIAAAWANMTQGSLFLPHFNSKTTFVTFVENGCARFEMATPYKFQRQQQOWPGQGQEE 366

Qy    544 EEDV--HYEQVRARLSKREAIVVLAGHPVFVSSGNENLLFAFGINAQNNHENFLAGRE 601
              : | : ||  |  | |:  | |  | :   | |  |||    | : :  | |  | |
Db    367 EEDMSENVHKVVSRVCKGEVFIVPAGHPFTILSQ-DQQFIAVGFGIYATNSKRTFLAG-E 424

Qy    602 RNVLQOIEPQAMELAFAAPRKEVEEESFNSQDQSIFFPGPROHQQQSPRSTKQQOPLVSIL 661
              ::|:  || |:|  |  : :| |   :|: :| |:| | |  ||| || || |:  | |||
Db    425 ENLLSNLNPAATRVTFGVGSKVAEKLFTSQNYSYFAPTSRS-QQQIPE--KHKPSFQSIL 481
```